Argumentation frameworks as Constraint Satisfaction Problems

Leila Amgoud • Caroline Devred

the date of receipt and acceptance should be inserted later

Abstract  Argumentation is a promising approach for defeasible reasoning. It consists of justifying each plausible conclusion by arguments. Since the available information may be inconsistent, a conclusion and its negation may both be justified. The arguments are thus said to be conflicting. The main issue is how to evaluate the arguments. Several semantics were proposed for that purpose. The most important ones are: stable, preferred, complete, grounded and admissible. A semantics is a set of criteria that should be satisfied by a set of arguments, called extension, in order to be acceptable. Different decision problems related to these semantics were defined (like whether an argumentation framework has a stable extension). It was also shown that most of these problems are computationally costly.

This paper studies how to encode the problem of computing the extensions of an argumentation framework (under each of the previous semantics) as a constraint satisfaction problem (CSP). Such encoding is of great importance since it makes it possible to use the very efficient solvers (developed by the CSP community) for computing the extensions. We focus on three families of argumentation frameworks: Dung’s abstract framework, its constrained version and preference-based argumentation framework.

Keywords  Default Reasoning, Argumentation, CSP

1 Introduction

Argumentation is a reasoning model based on the construction and evaluation of interacting arguments. An argument is a reason for believing in a statement, doing an action, pursuing a goal, etc. Argumentation theory is gaining an increasing interest in Artificial Intelligence, namely for reasoning about defeasible/uncertain information, making decisions under uncertainty, learning concepts, and modeling agents’ interactions (see [1] for a survey of the various works done on argumentation in AI and [29,

Leila Amgoud
IRIT - CNRS
118, route de Narbonne, 31062 Toulouse Cedex 9, France
E-mail: amgoud@irit.fr
5,21] for some works on complexity issues).

The most abstract argumentation framework was proposed in the seminal paper [19] by Dung. It consists of a set of (equally important) arguments, a binary relation representing attacks among arguments, and semantics for evaluating the arguments. A semantics describes when a set of arguments, called extension, is acceptable without bothering neither on the origin nor on the structure of the arguments/attacks. The most important semantics are grounded (which ensures only one extension), admissible, complete, stable and preferred semantics which may lead to more than one extension. It is worth mentioning that generally an argumentation framework is represented by a graph whose nodes are the arguments of the framework and the edges are the attacks among arguments.

Dung’s framework was extended in different ways in the literature. In [2, 4], arguments are assumed to not have the same strength. Thus, in addition to the attack relation, another binary relation on the set of arguments is available. It represents preferences between arguments. In [13], arguments are assumed to have the same strength, but a constraint on arguments may be available. This constraint represents a desirable property that should be satisfied by the extensions, i.e., by the acceptable sets of arguments. It is worth mentioning that in both works, Dung’s semantics are used to evaluate arguments, thus to compute the extensions.

Different decision problems related to the implementation of the previous semantics were identified (like for instance testing whether a framework has a stable extension, or whether an argument belongs to one (or every) extension, or whether a given set of arguments is an extension under a given semantics). The computational complexity of each identified problem was studied (see for instance [13, 17, 18, 23]). The results are disappointing since they show that the most important decision problems are very costly. For instance, checking whether a set of arguments is a preferred extension of an argumentation framework is CO-NP-complete, and checking whether a framework has any stable extension is NP-complete. Some algorithms that compute extensions under some semantics were developed, for instance in [11, 16, 24]. However, the efficiency of those algorithms was not proved. They are neither tested on benchmarks nor compared to other algorithms developed for the same purpose.

In [6], Besnard and Doutre defined Dung’s semantics as satisfiability problems. Indeed, for each semantics, the argumentation framework is translated into a set of propositional formulas. The models of this set correspond exactly to the extensions of the framework under the chosen semantics. This work is of great importance since it makes it possible to use the efficient SAT solvers for computing the extensions of argumentation frameworks.

Besides, there is a huge literature on Constraints Satisfaction Problems (CSP) since many real-world problems can be described as CSPs (e.g., [12, 33, 15, 26, 28, 30, 32]). A CSP consists of a set of variables, a (generally finite) domain for each variable and a set of constraints. Each constraint is defined over a subset of variables and limits the combination of values that the variables in this subset can take. The goal is to find an assignment to the variables which satisfies all the constraints. In some problems, the goal is to find all such assignments. Solving a constraint satisfaction problem on a finite domain is an NP-complete problem in general. In order to be solved in a reasonable time, different solvers were developed. They use a form of search based on variants of backtracking, constraint propagation and local search [28].
CSP and SAT are considered as two independent threads of research. They have a lot in common as evidenced by similar ideas underlying the branch and prune algorithms that are most successful at solving both kinds of problems. They also exhibit differences in the way they are used to state and solve problems (see [8] for a survey of the similarities and differences between CSP and SAT). It is well known that CSP is more expressive for the modeling phase. In an argumentation context, CSP is more appropriate for encoding complex semantics like grounded and preferred. It is also possible for users to add constraints as in [13] into CSP while in SAT there is usually little room for this parametrization.

In this paper, we extend the work of Besnard and Doutre [6] in three ways: first we study all Dung’s semantics including preferred and grounded semantics which were not considered in [6]. Second, we study the semantics in three families of frameworks: Dung’s framework [19] as in [6] but also in constrained argumentation framework [13] and preference-based argumentation framework [2]. Finally, we encode the various semantics as CSPs. We show that each semantics is captured by one particular CSP. In each CSP, the arguments of the framework play the role of variables whose domains are all binary. For those semantics that are studied in [6] we take advantage of the possible translation from SAT to CSP in order to define their corresponding CSPs.

This paper is organized as follows: Section 2 recalls the basic concepts of a CSP. Section 3 recalls Dung’s argumentation framework and shows how to compute its extensions (under the studied semantics) by CSPs. Section 4 presents the constrained version of Dung’s framework and proposes different CSPs that compute the extensions under the same semantics of such a framework. Section 5 recalls the last family of argumentation frameworks (preference-based argumentation frameworks) as well as their encoding as CSPs. In section 6, we compare our approach to existing works on the topic. The last section is devoted to concluding remarks and perspectives.

2 Constraint satisfaction problems (CSPs)

A constraint satisfaction problem (or CSP) is expressed as follows: a set of variables, finite sets of possible values that can be assigned to the variables (called domains), and a list of constraints. Each constraint involves some subset of the variables and specifies the allowable combinations of values for that subset. The idea is to find the possible combinations of values of the variables that satisfy every constraint. It is worth mentioning that many real-world problems such as scheduling fall into this framework (see [30] for some applications). Another example of the CSP is the timetabling problem in which courses are given to different classes, in different rooms, and by different lecturers. The same lecturer cannot give two courses at the same time, and two courses cannot take place at the same time, etc.

Definition 1 (CSP) A CSP instance is a triple \((\mathcal{X}, \mathcal{D}, \mathcal{C})\) where:

- \(\mathcal{X} = \{x_1, \ldots, x_n\}\) is a set of variables,
- \(\mathcal{D} = \{D_1, \ldots, D_n\}\) where \(D_i\) is a finite domain for the variable \(x_i \in \mathcal{X}\), and
- \(\mathcal{C} = \{c_1, \ldots, c_m\}\) is a set of constraints.

Each constraint \(c_i\) is a pair \((h_i, R_i)\) where...
\[- h_i = (x_{i1}, \ldots, x_{ik}) \text{ where } x_{ij} \in X \]
\[- H_i \subseteq D_{i1} \times \ldots \times D_{ik} \text{ where } D_{ij} \text{ is the domain of } x_{ij} \in h_i.\]

Note that \( H_i \) is a \( k \)-ary relation over \( D \). It is a subset of all allowed combinations of values for the variables in \( h_i \). A state of the problem is defined by an assignment of values to some or all of the variables.

**Definition 2 (Assignment)** An assignment \( v \) for a CSP instance \((X, D, C)\) is a mapping that assigns to every variable \( x_i \in X \) an element \( v(x_i) \in D_i \). An assignment \( v \) satisfies a constraint \((x_{i1}, \ldots, x_{ik}), H_i) \in C \) iff \((v(x_{i1}), \ldots, v(x_{ik})) \in H_i.\)

Finally, a solution of a CSP is an assignment that satisfies its constraints.

**Definition 3 (Solution)** A solution of a CSP instance \((X, D, C)\) is an assignment \( v \) that satisfies all the constraints in \( C \) and in which all the variables of \( X \) are assigned a value. We write \((v(x_1), \ldots, v(x_n))\) to denote the solution.

### 3 Abstract argumentation framework

This section recalls Dung’s argumentation framework and presents the different corresponding CSPs which return its extensions under various semantics.

#### 3.1 Dung’s framework

In [19], Dung has developed the most abstract argumentation framework in the literature. It consists of a set of arguments and an attack relation between them.

**Definition 4 (Argumentation framework)** An argumentation framework \((\mathcal{F}, R)\) is a pair \(\mathcal{F} = (\mathcal{A}, R)\) where \(\mathcal{A}\) is a set of arguments and \(R\) is an attack relation \((R \subseteq \mathcal{A} \times \mathcal{A})\). The notations \(aRb\) or \((a, b) \in R\) mean that the argument \(a\) attacks the argument \(b\).

Different acceptability semantics for evaluating arguments were proposed in the same paper [19]. A semantics is a set of criteria that should be satisfied by a set of arguments, called extension, in order to be acceptable. Some semantics ensure only one extension to every argumentation framework whereas others may lead to several extensions. Dung’s semantics have in common two basic properties: conflict-freeness and defence.

**Definition 5** Let \(\mathcal{F} = (\mathcal{A}, R)\) be an AF, \(a \in \mathcal{A}\) and \(B \subseteq \mathcal{A}\).

- \(B\) is conflict-free iff \(\not\exists a, b \in B \text{ s.t. } aRb.\)
- \(B\) defends an argument \(a\) iff for all \(b \in \mathcal{A} \text{ s.t. } bRa\), there exists \(c \in B \text{ s.t. } cRb.\)

The following definition recalls all the acceptability semantics proposed in [19].

**Definition 6 (Acceptability semantics)** Let \(\mathcal{F} = (\mathcal{A}, R)\) be an AF and \(B \subseteq \mathcal{A}\).

- \(B\) is an admissible set iff it is conflict-free and defends its elements
- \(B\) is a preferred extension iff it is a maximal (for set \(\subseteq\)) admissible set.
- \(B\) is a stable extension iff it is conflict-free and attacks any argument in \(\mathcal{A} \setminus B.\)
- $\mathcal{B}$ is a complete extension iff it is conflict-free and contains all arguments it defends.
- $\mathcal{B}$ is a grounded extension iff it is a minimal (for set $\subseteq$) complete extension.

It was shown in [19] that any argumentation framework has a unique grounded extension (which may be empty). It may have several admissible, complete, stable and preferred extensions. However, unlike other semantics, stable extensions do not always exist. Finally, it can be checked that any stable extension is a preferred one but the reverse is not always true.

**Example 1.** Let us consider the framework $\mathcal{F}_1 = (\mathcal{A}_1, \mathcal{R}_1)$ where $\mathcal{A}_1 = \{a, b, c, d\}$ and let $\mathcal{R}_1$ be as depicted in the figure below.

![Diagram](image)

The framework $\mathcal{F}_1$ has two preferred extensions which are also stable: $\mathcal{B}_1 = \{a, c\}$ and $\mathcal{B}_2 = \{b, d\}$. The sets $\mathcal{B}_1$, $\mathcal{B}_2$ and $\mathcal{B}_3 = \{\}$ are the admissible and complete extensions of the framework, whereas $\mathcal{B}_3$ is its grounded extension.

### 3.2 Computing Dung’s semantics by CSPs

In this section, we propose different mappings of Dung’s argumentation framework into CSP instances. The idea is: starting from an argumentation framework, we define a CSP instance whose solutions are the extensions of the framework under a given acceptability semantics. In all the instances, arguments play the role of variables that is, a variable is associated to each argument. Each variable may take two values 0 or 1 meaning that the corresponding argument is rejected or accepted. Thus, the domains of the variables are all binary. Things are different with the constraints. We show that according to the semantics that is studied, the definition of a constraint changes.

Let us start with a CSP instance that computes the conflict-free sets of arguments. The idea here is that a set of arguments cannot contain two conflicting arguments. Thus, each attack $(a, b) \in \mathcal{R}$ gives birth to a constraint which says that the two variables $a$ and $b$ cannot take value 1 at the same time. This constraint has the following form: $((a, b), ((0, 0), (0, 1), (1, 0)))$. The combination $(0, 0)$ means that the two arguments do not belong to the set. It is also worth noticing that the combinations $(0, 0), (0, 1), (1, 0)$ are exactly the models of the propositional formula $a \Rightarrow \neg b$ (or $b \Rightarrow \neg a$). Both formulas $a \Rightarrow \neg b$ and $b \Rightarrow \neg a$ can be used since in conflict-freeness, the orientation of the attack relation is not important.

**Notations:** Throughout the paper, $|x|$ denotes the models of a propositional formula $x$, and $\text{Atoms}(x)$ the set of atoms that are involved in $x$.

**Definition 7** (Free CSP) Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation framework. A free CSP associated with $\mathcal{F}$ is a tuple $(\mathcal{X}, \mathcal{D}, \mathcal{C})$ where $\mathcal{X} = \mathcal{A}$, $\forall \mathcal{D}_i \in \mathcal{D}$, $\mathcal{D}_i = \{0, 1\}$ and $\mathcal{C} = \{(h, H) | h = (a, b) \text{ where } (a, b) \in \mathcal{R} \text{ and } H = [a \Rightarrow \neg b]\}$. 
It can be checked that there are as many constraints as attacks in the argumentation framework.

**Property 1** Let \((X, D, C)\) be the free CSP instance associated with the AF \(F = (A, R)\). It holds that \(|C| = |R|\).

**Proof** This follows immediately from Definition 7.

The following result shows that the solutions of this CSP are the conflict-free sets of arguments of the corresponding AF.

**Theorem 1** Let \((X, D, C)\) be the free CSP instance associated with the AF \(F = (A, R)\). The tuple \((v(x_1), \ldots, v(x_n))\) is a solution of the CSP iff the set \(\{x_1, \ldots, x_k\} \subseteq X\) s.t. \(v(x_i) = 1\) is conflict-free (with \(i = j, \ldots, k\)).

**Proof** Let \((X, D, C)\) be the free CSP associated with the argumentation framework \(F = (X, R)\). Assume that \((v(x_1), \ldots, v(x_n))\) is a solution of the CSP. Let \(X = \{x_j, \ldots, x_k\}\) be such that \(v(x_j) = 1\). Assume also that \(X\) is not conflict-free, thus \(\exists x, x' \in X\) such that \(xRx'\). Thus, \(\exists c \in C\) such that \(c = ((x, x'), \{x = \neg x'\})\). Since \((v(x_1), \ldots, v(x_n))\) is a solution, then it satisfies the constraint \(c\). However, \(v(x_1) = 1\) and \(v(x') = 1\), then \((1, 1) \notin \{x = \neg x'\}\). Contradiction. Assume now a tuple \((v(x_1), \ldots, v(x_n))\) such that \(X = \{x_j, \ldots, x_k\}\) (with \(v(x_j) = 1\)) is conflict-free. Assume that \((v(x_1), \ldots, v(x_n))\) is not a solution. Thus, there exists a constraint \(c = ((x, x'), \{x = \neg x'\})\) which is violated. But since \(X\) is conflict-free, then from the definition of Free CSP, this constraint cannot exist between \(x\) and \(x'\).

Let us illustrate this kind of CSP by considering the argumentation framework \(F_1\) given in Example 1.

**Example 1 (Cont.):** The free CSP corresponding to \(F_1\) is \((X, D, C)\) s.t.

- \(X = \{a, b, c, d\}\),
- \(D = \{\{0, 1\}, \{0, 1\}, \{0, 1\}\}\),
- \(C = \{((a, b), [a \Rightarrow \neg b]), ((b, c), [b \Rightarrow \neg c]), ((c, d), [c \Rightarrow \neg d]), ((d, a), [d \Rightarrow \neg a])\}\).

This CSP has the following solutions:

<table>
<thead>
<tr>
<th></th>
<th>((0, 0, 0, 0))</th>
<th>((1, 0, 0, 0))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>((0, 1, 0, 0))</td>
<td>((0, 0, 1, 0))</td>
</tr>
<tr>
<td></td>
<td>((0, 0, 1, 1))</td>
<td>((1, 0, 1, 0))</td>
</tr>
<tr>
<td></td>
<td>((0, 1, 0, 1))</td>
<td></td>
</tr>
</tbody>
</table>

Thus, the sets \(\{\}, \{a\}, \{b\}, \{c\}, \{d\}, \{a, c\}\) and \(\{b, d\}\) are conflict-free.

Let us now study the case of stable semantics. Stable extensions are computed by a CSP which considers that an argument and its attackers cannot have the same value.

**Definition 8 (Stable CSP)** Let \(F = (A, R)\) be an argumentation framework. A stable CSP associated with \(F\) is a tuple \((X, D, C)\) where \(X = A\), \(\forall D_1 \in D, D_1 = \{0, 1\}\) and \(C = \{(h, H) \mid h = (a_1, b_1, \ldots, b_0)\) s.t. \((b_1, a) \in R\) and \(H = \{a \Rightarrow \neg b_1, a \in A\}\), with the convention that \(\bigwedge_{x \in X} x = \top\) if \(X = \emptyset\).
Unlike free CSP, the number of constraints is equal to the number of arguments of the framework.

Property 2 Let \((X, D, C)\) be the stable CSP instance associated with the AF \(F = (A, R)\). It holds that \(|C| = |A|\).

Proof This follows immediately from Definition 8.

It is worth mentioning that the previous definition is inspired from [14]. The following result shows the correspondence between the solutions of a stable CSP and the stable extensions of its associated argumentation framework.

Theorem 2 Let \((X, D, C)\) be a stable CSP associated with \(F = (A, R)\). The tuple \((v(x_1), \ldots, v(x_n))\) is a solution of the CSP if and only if the set \(\{x_j, \ldots, x_k\}\) s.t. \(v(x_i) = 1\) is a stable extension of \(F\).

Proof Let \((X, D, C)\) be a stable CSP associated with \(F = (A, R)\). In [6], it was shown that a set \(E \subseteq A\) is a stable extension of the argumentation framework \(F\) if \(E\) is a model of the formula \(\bigwedge_{a \in A} (a \iff \bigwedge_{b \in \{h \mid a \in R\}} \neg b)\), thus a model of the set \(\{a \iff \bigwedge_{b \in \{h \mid a \in R\}} \neg b, a \in A\}\). Consequently, \(E\) is a stable extension if and only if it satisfies all the constraints of \(C\), thus if and only if it is a solution of the stable CSP.

Let us illustrate the notion of stable CSP on our running Example 1.

Example 1 (Cont): The stable CSP corresponding to \(F_1\) is \((X, D, C)\) s.t. \(X = \{a, b, c, d\}\), \(D = \{\{0, 1\}, \{0, 1\}, \{0, 1\}\}\), and \(C = \{\{(a, d), \{a \iff \neg d\}\}, \{(b, a), \{b \iff \neg a\}\}, \{(c, b), \{c \iff \neg b\}\}, \{(d, c), \{d \iff \neg c\}\}\}\). This CSP has two solutions: \((1, 0, 1, 0)\) and \((0, 1, 0, 1)\). The sets \(\{a, c\}\) and \(\{b, d\}\) are the two stable extensions of \(F_1\).

Let us now consider the case of an argumentation framework which has no stable extensions.

Example 2 Let \(F_2 = (A_2, R_2)\) be an argumentation framework where \(A_2 = \{a, b\}\) and \(R_2 = \{(b, b)\}\). Note that this framework has no stable extension. The corresponding stable CSP is \((X, D, C)\) where \(X = A_2\) and \(C = \{\{(a), \{a\}\}, \{(b), \{b \land \neg b\}\}\}\). This CSP has no solution since the constraint \(b \land \neg b\) is not satisfiable. Consequently, \(F_2\) has no stable extensions.

The two previous CSPs are simple since attacks are directly transformed into constraints. The notion of defence is not needed in both cases. However, things are not so obvious with admissible semantics. Indeed, a CSP that computes admissible extensions contains, for each argument, two kinds of constraints: constraints that ensure that this argument cannot be simultaneously accepted with any of its attackers, and constraints that ensure that the argument is defended by other arguments. Note that the first kind of constraints ensures conflict-freeness whereas the second ensures the defence requirement of the semantics.

Definition 9 (Admissible CSP) Let \(F = (A, R)\) be an argumentation framework. An admissible CSP associated with \(F\) is a tuple \((X, D, C)\) where \(X = A\), \(\forall D_i \in D, D_i = \{0, 1\}\) and \(C = C_1 \cup C_2\) where
\[ C_1 = \{(h, H) \mid h = (a, b_1, \ldots, b_m), H = \bigwedge_{(b, a) \in R} \neg b_1 \mid a \in A\} \]
\[ C_2 = \{(h, H) \mid h = (a, c_1, \ldots, c_n), H = \bigwedge_{b(b, a) \in R} \bigvee_{(c, b) \in R} c_i \mid a \in A\} \]

It is worth mentioning that the two kinds of constraints can be combined in a unique class, where each constraint is a conjunction of a formula in \(C_1\) and a formula of \(C_2\) for the same argument. However, such encoding would not be optimal. It makes Generalized Arc Consistency possible but not polynomial [31].

The number of constraints is at most the number of arguments of the CAF plus one (the constraint of the CAF).

**Property 3** Let \((X, D, C)\) be the admissible CSP instance associated with the AF \(F = (A, R)\). It holds that \(|C| \leq 2 \times |A|\).

**Proof** This follows immediately from the previous definition.

The following result shows that the solutions of an admissible CSP provide the admissible extensions of the corresponding argumentation framework.

**Theorem 3** Let \((X, D, C)\) be an admissible CSP associated with an AF \(F\). The tuple \((v(x_1), \ldots, v(x_n))\) is a solution of this CSP iff the set \(\{x_1, \ldots, x_k\}\) s.t. \(v(x_i) = 1\) is an admissible set of \(F\).

**Proof** Let \((X, D, C)\) be an admissible CSP associated with an AF \(F\). In [6], it was shown that a set \(E \subseteq A\) is an admissible extension of the argumentation framework \(F\) iff \(E\) is a model of the formula \(\bigwedge_{a \in A} (a \Rightarrow \bigwedge_{(b, a) \in R} \neg b_1) \land (a \Rightarrow \bigwedge_{b(b, a) \in R} \bigvee_{(c, b) \in R} c_i)\).

Consequently, \(E\) is an admissible extension iff it satisfies all the constraints of \(C\), thus iff it is a solution of the admissible CSP.

Let us illustrate the notion of admissible CSP with a simple example.

**Example 3** Let us consider the framework \(F_3 = (A_3, R_3)\) where \(A_3 = \{a, b, c, d\}\) and let \(R_3\) be as depicted in figure below:

![Graph](image)

The admissible CSP associated with \(F_3\) is \((X, D, C)\) where: \(X = A_3, D = \{\{0, 1\}, \{0, 1\}, \{0, 1\}\} \) and \(C = \{(d), |d \Rightarrow \top| \}, \{(c), |c \Rightarrow \top| \}, \{(b, c, d), |b \Rightarrow \neg c \land \neg d| \}, \{(a, b), |a \Rightarrow \neg b| \}, \{(a, c, d), |a \Rightarrow c \lor d| \} \).

This CSP has the following solutions: \((0, 0, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1), (0, 0, 1, 1)\) \((1, 0, 1, 0), (1, 0, 0, 1), (1, 0, 1, 1)\). These solutions return the admissible sets of \(F_3\), that is: \(\emptyset, \{c\}, \{d\}, \{c, d\}, \{a, c\}, \{a, d\}\) and \(\{a, c, d\}\).
As preferred extensions are maximal (for set inclusion) admissible sets, then they are computed by an admissible CSP.

**Theorem 4** Let \((X, \mathcal{D}, C)\) be an admissible CSP associated with an AF \(\mathcal{F}\). Each maximal (for set inclusion) set \(\{x_j, \ldots, x_k\}\), s.t. \(v(x_i) = 1\) and \((v(x_1), \ldots, v(x_n))\) is a solution of the CSP, is a preferred extension of \(\mathcal{F}\).

**Proof** This follows directly from the definition of a preferred extension and Theorem 3.

Let us now come back to Example 3 and check the preferred extensions of the framework.

**Example 3 (Cont):** The solution \((1, 0, 1, 1)\) returns the only preferred extension of \(\mathcal{F}_3\), i.e. \(\{a, c, d\}\).

Complete extensions are also computed by a CSP which takes into account the notion of defence in the constraints.

**Definition 10 (Complete CSP)** Let \(\mathcal{F} = (\mathcal{A}, \mathcal{R})\) be an argumentation framework. A complete CSP associated with \(\mathcal{F}\) is a tuple \((X, D, C)\) where \(X = \mathcal{A}\), for each \(D_i \in D, D_i = \{0, 1\}\) and \(C = C_1 \cup C_2\) where

\[
- C_1 = \{(h, H) \mid h = (a, b_1, \ldots, b_m), H = \bigwedge_{(b_i, a) \in R} \neg b_i \mid a \in \mathcal{A}\}
\]

\[
- C_2 = \{(h, H) \mid h = (a, c_1, \ldots, c_n), H = \bigwedge_{b \in \mathcal{R}} \bigwedge_{c_i \in \mathcal{R}} (\bigvee c_{i \in R}) \mid a \in \mathcal{A}\}
\]

Note that there is a slight difference between the constraints of an admissible CSP and those of a complete CSP. Since a complete extension should contain all the arguments it defends, then an argument and all its defenders should be in the same set. However, the only requirement on an admissible set is that it defends its arguments. This is encoded by a simple logical implication.

**Property 4** Let \((X, D, C)\) be the complete CSP instance associated with the AF \(\mathcal{F} = (\mathcal{A}, \mathcal{R})\). It holds that \(|C| \leq 2 \cdot |A|\).

**Proof** This follows immediately from the previous definition.

The next result confirms that a complete CSP returns all the complete extensions of an AF.

**Theorem 5** Let \((X, D, C)\) be a complete CSP associated with an AF \(\mathcal{F}\). The tuple \((v(x_1), \ldots, v(x_n))\) is a solution of the CSP iff the set \(\{x_j, \ldots, x_k\}\), s.t. \(v(x_i) = 1\) is a complete extension of \(\mathcal{F}\).

**Proof** Let \((X, D, C)\) be a complete CSP associated with an AF \(\mathcal{F}\). In [6], it was shown that a set \(E \subseteq \mathcal{A}\) is a complete extension of the argumentation framework \(\mathcal{F}\) iff \(E\) is a model of the formula \(\bigwedge_{a \in \mathcal{A}} (a \Rightarrow \bigwedge_{(b, a) \in \mathcal{R}} \neg b_i) \wedge (a \Leftrightarrow \bigwedge_{b \in \mathcal{R}} \bigwedge_{c_i \in \mathcal{R}} (\bigvee c_{i \in R}))\). Consequently, \(E\) is a complete extension iff it satisfies all the constraints of \(C\), thus iff it is a solution of the complete CSP.
Example 3 (Cont): The complete CSP associated with $\mathcal{F}_3$ is $(X, D, C)$ where: $X = A_3$, $D = \{\{0, 1\}, \{0, 1\}, \{0, 1\} \}$ and $C =$

$\{((d), |d \Rightarrow \top|),$
$((c), |c \Rightarrow \top|),$
$((b, c, d), |b \Rightarrow \neg c \land \neg d|),$
$((a, b), |a \Rightarrow \neg b|),$
$((a, c, d), |a \Leftrightarrow c \lor \neg d|)\}.$

This CSP has one solution which is $(1, 0, 1, 1)$. Thus, the set $\{a, c, d\}$ is the unique complete extension of $\mathcal{F}_3$.

Since grounded extension is a minimal (for set inclusion) complete extension, then it is computed by a complete CSP as follows

**Theorem 6** Let $(X, D, C)$ be a complete CSP associated with an AF $\mathcal{F}$. The grounded extension of $\mathcal{F}$ is the minimal (for set inclusion) set $\{x_1, \ldots, x_k\}$ s.t. $v(x_i) = 1$ and $(v(x_1), \ldots, v(x_n))$ is a solution of the CSP.

**Proof** This follows from the definition of a grounded extension and Theorem 5.

Example 3 (Cont): The grounded extension of $\mathcal{F}_3$ is $\{a, c, d\}$ which is returned by the unique solution of the complete CSP corresponding to $\mathcal{F}_3$.

4 Constrained framework

This section recalls the constrained version of Dung’s argumentation framework [13] and proposes various CSPs that compute its extensions under Dung’s semantics.

4.1 Basic definitions

The basic argumentation framework of Dung was extended in [13] by adding a constraint on arguments. This constraint should be satisfied by Dung’s extensions (under a given semantics). For instance, in Example 1, one may imagine a constraint which requires that the two arguments $a$ and $c$ belong to the same extension. Note that this constraint is satisfied by $B_1$ but not by $B_2$. Thus, $B_1$ would be the only extension of the framework. The constrained version of Dung’s system may be useful in some AI problems, like practical reasoning [3], where a filter needs to be applied on the extensions of the argumentation framework in order to select only the ones that contain for each desire, an argument justifying the desire and another argument ensuring its feasibility.

In this framework, the constraint is a formula of a propositional language $\mathcal{L}_A$ whose alphabet is exactly the set $A$ of arguments. Thus, each argument in $A$ is a literal of $\mathcal{L}_A$. $\mathcal{L}_A$ contains all the formulas that can be built using the usual logical operators ($\land, \lor, \Rightarrow, \neg, \Leftrightarrow$) and the constant symbols ($\top$ and $\bot$).

**Definition 11 (Constraint, Completion)** Let $A$ be a set of arguments and $\mathcal{L}_A$ its corresponding propositional language.
- $C$ is a constraint on $A$ if $C$ is a formula of $\mathcal{L}_A$.
- The completion of a set $B \subseteq A$ is $\bar{B} = \{a \mid a \in B\} \cup \{-a \mid a \in A \setminus B\}$.
- A set $B \subseteq A$ satisfies $C$ if $\bar{B}$ is a model of $C$ ($\bar{B} \models C$).

The completion of a set $B$ of arguments is a set in which each argument of $A$ appears either as a positive literal if the argument belongs to $B$ or as a negative one otherwise. Thus, $|\bar{B}| = |A|$. A constrained argumentation framework (CAF) is defined as follows:

**Definition 12 (CAF)** A constrained argumentation framework (CAF) is a triple $F = (A, R, C)$ where $A$ is a set of arguments, $R \subseteq A \times A$ is an attack relation and $C$ is a constraint on the set $A$.

Dung’s semantics are extended to the case of CAFs. The idea is to compute Dung’s extensions (under a given semantics), and to keep among those extensions only the ones that satisfy the constraint $C$.

**Definition 13 (C-admissible set)** Let $F = (A, R, C)$ be a CAF and $B \subseteq A$. The set $B$ is $C$-admissible in $F$ iff:

1. $B$ is admissible,
2. $B$ satisfies the constraint $C$.

In [19], it was shown that the empty set is always admissible. However, it is not always $C$-admissible since the set $\emptyset$ does not always imply $C$.

**Definition 14 (C-preferred, C-stable extension)** Let $F = (A, R, C)$ be a CAF and $B \subseteq A$.

- $B$ is a $C$-preferred extension of $F$ if $B$ is maximal for set-inclusion among the $C$-admissible sets.
- $B$ is a $C$-stable extension of $F$ if $B$ is a $C$-preferred extension that attacks all arguments in $A \setminus B$.

The following result summarizes the links between the extensions of a CAF $F = (A, R, C)$ and those of its basic version $F' = (A, R')$ (i.e. the argumentation framework without the constraint).

**Theorem 7** [13] Let $F = (A, R, C)$ be a CAF and $F' = (A, R)$ be its basic version.

- For each $C$-preferred extension $B$ of $F$, there exists a preferred extension $B'$ of $F'$ s.t. $R \subseteq B'$
- Every $C$-stable extension of $F$ is a stable extension of $F'$. The converse does not hold

It is worth noticing that when the constraint of a CAF is a tautology, then the extensions of this CAF coincide with those of its basic version.

Let us now illustrate this notion of CAFs through a simple example

**Example 1 (Cont):** Assume an extended version of the argumentation framework $F_1$ where we would like to accept the two arguments $a$ and $c$. This is encoded by a constraint $C : a \land c$. It can be checked that the CAF $(A_1, R_1, C)$ has one $C$-stable extension which is $B_1 = \{a, c\}$. Note that $B_2 = \{b, d\}$ is a stable extension of $F_1$ but not a $C$-stable extension of its constrained version.
4.2 Mappings into CSPs

Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}, \mathcal{C})$ be a given CAF. In order to compute its $\mathcal{C}$-extensions under different semantics, we follow the same line of research as in the previous section. The only difference is that in addition to the constraints defined in Section 3.2, there is an additional constraint which is $\mathcal{C}$.

Let us start with $\mathcal{C}$-stable extensions. They are computed by the stable CSP given in Def. 8 augmented by the constraint $\mathcal{C}$ in its set $\mathcal{C}$.

**Definition 15 ($\mathcal{C}$-stable CSP)** Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}, \mathcal{C})$ be a constrained argumentation framework. A $\mathcal{C}$-stable CSP associated with $\mathcal{F}$ is a tuple $\langle \mathcal{X}, \mathcal{D}, \mathcal{C} \rangle$ where $\mathcal{X} = \mathcal{A}$, $\forall \mathcal{D}_i \in \mathcal{D}$, $\mathcal{D}_i = \{0, 1\}$ and $\mathcal{C} = \{(h, H) \mid h = (a, b_1, \ldots, b_n) \text{ s.t. } (b_i, a) \in \mathcal{R} \text{ and } H = [a \Leftrightarrow \bigwedge \neg b_i], \ a \in \mathcal{A}\} \cup \{(h', H') \mid H' = [\mathcal{C}], h' = (a_1, \ldots, a_j) \text{ where } \{a_1, \ldots, a_j\} = \text{Atoms}(\mathcal{C})\}$.

It is easy to check that the number of constraints is exactly the number of arguments of the CAF plus one (the constraint of the CAF).

**Property 5** Let $\langle \mathcal{X}, \mathcal{D}, \mathcal{C} \rangle$ be the $\mathcal{C}$-stable CSP instance associated with the CAF $\mathcal{F} = (\mathcal{A}, \mathcal{R}, \mathcal{C})$. It holds that $|\mathcal{C}| = |\mathcal{A}| + 1$.

**Proof** This follows immediately from the previous definition.

We show next that the solutions of a $\mathcal{C}$-stable CSP return all the $\mathcal{C}$-stable extensions of the corresponding CAF.

**Theorem 8** Let $\langle \mathcal{X}, \mathcal{D}, \mathcal{C} \rangle$ be a $\mathcal{C}$-stable CSP associated with a CAF $\mathcal{F} = (\mathcal{A}, \mathcal{R}, \mathcal{C})$. The tuple $(v(x_1), \ldots, v(x_k))$ is a solution of the CSP iff the set $\{x_1, \ldots, x_k\}$ such that $v(x_i) = 1$ is a $\mathcal{C}$-stable extension of $\mathcal{F}$.

**Proof** Let $\mathcal{T} = (\mathcal{X}, \mathcal{D}, \mathcal{C}')$ be a $\mathcal{C}$-stable CSP associated with a CAF $(\mathcal{A}, \mathcal{R}, \mathcal{C})$. The $\mathcal{C}$-stable extensions of the CAF $(\mathcal{A}, \mathcal{R}, \mathcal{C})$ are the stable extensions of $(\mathcal{A}, \mathcal{R})$ that verify the constraint $\mathcal{C}$. Besides, let $\mathcal{T}' = (\mathcal{X}, \mathcal{D}, \mathcal{C}')$ be the stable CSP associated with an AF $(\mathcal{A}, \mathcal{R})$. From Theorem 2, there is a bijection between the solutions of the CSP $\mathcal{T}'$ and the stable extensions of $(\mathcal{A}, \mathcal{R})$. Moreover, from Definition 15, the solutions of $\mathcal{T}$ verify all the constraints in $\mathcal{C}'$ and the additional constraint $\mathcal{C}$. Thus, there is a bijection between the solutions of the $\mathcal{C}$-stable CSP and the $\mathcal{C}$-stable extensions of $(\mathcal{A}, \mathcal{R}, \mathcal{C})$.

**Example 1 (Cont.):** The $\mathcal{C}$-stable CSP associated with the CAF extending $\mathcal{F}_1$ with the constraint $\mathcal{C}$: $a \land c$ is $(\mathcal{X}, \mathcal{D}, \mathcal{C})$ s.t. $\mathcal{X} = \{a, b, c, d\}$, $\mathcal{D} = \{\{0, 1\}, \{0, 1\}, \{0, 1\}\}$, and $\mathcal{C} = \{(a, c) \land c\}, \{(a, d) \leftarrow \neg d\}, \{(b, a), \{b \leftarrow \neg a\}, \{(c, b), \{c \leftarrow \neg b\}, \{(d, c), \{d \leftarrow \neg c\}\}$. This CSP has one solution which is $(1, 0, 1, 0)$. It returns the $\mathcal{C}$-stable extension $(a, c)$ of the CAF.

A CSP which computes the $\mathcal{C}$-admissible sets of a CAF is grounded on the admissible CSP introduced in Definition 9.

**Definition 16 ($\mathcal{C}$-admissible CSP)** Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}, \mathcal{C})$ be a constrained argumentation framework. A $\mathcal{C}$-admissible CSP associated with $\mathcal{F}$ is a tuple $\langle \mathcal{X}, \mathcal{D}, \mathcal{C} \rangle$ where $\mathcal{X} = \mathcal{A}$, for each $a_i \in \mathcal{X}$, $\mathcal{D}_i = \{0, 1\}$ and $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3$ where
\( C_1 = \{ (h, H) \mid h = (a, b_1, \ldots, b_m), H = \bigwedge_{(b, a) \in R} \neg b_i | a \in A \} \)

\( C_2 = \{ (h, H) \mid h = (a, c_1, \ldots, c_n), H = \bigvee_{b; (b, a) \in R} \bigvee_{(c, b) \in R} c_i | a \in A \} \)

\( C_3 = \{ (h', H') \mid H' = [\mathcal{C}], h' = (a_1, \ldots, a_j) \text{ where } \{a_1, \ldots, a_j\} = \text{Atoms}(\mathcal{C}) \} \)

Property 6 Let \((X, \mathcal{D}, \mathcal{C})\) be the \mathcal{C}-admissible CSP instance associated with the CAF \(F = (A, R, \mathcal{C})\). It holds that \(|\mathcal{C}| = 2 \cdot |A| + 1\).

Proof This follows immediately from the previous definition.

Theorem 9 Let \((X, \mathcal{D}, \mathcal{C})\) be a \mathcal{C}'-admissible CSP associated with a CAF \(F = (A, R, \mathcal{C})\). The tuple \((v(x_1), \ldots, v(x_n))\) is a solution of the CSP iff the set \(\{x_1, \ldots, x_k\}\) s.t. \(v(x_1) = 1\) is a \mathcal{C}'-admissible set of the CAF \(F\).

Proof Let \(T = (X, \mathcal{D}, \mathcal{C})\) be a \mathcal{C}'-admissible CSP associated with a CAF \((A, R, \mathcal{C})\). The \mathcal{C}'-admissible extensions of the CAF \((A, R, \mathcal{C})\) are the admissible extensions of \((A, R)\) that verify the constraint \mathcal{C}'. Besides, let \(T' = (X, \mathcal{D}, \mathcal{C}')\) be the admissible CSP associated with a AF \((A, R)\). From Theorem 3, there is a bijection between the solutions of the CSP \(T\) and the admissible extensions of \((A, R)\). Moreover, from Definition 16, the solutions of \(T\) verify all the constraints in \mathcal{C}' and the additional constraint \mathcal{C}'. Thus, there is a bijection between the solutions of the \mathcal{C}'-admissible CSP and the \mathcal{C}'-admissible extensions of \((A, R, \mathcal{C}')\).

What about the \mathcal{C}'-preferred extensions of a constrained argumentation framework? Recall that \mathcal{C}'-preferred extensions are maximal (for set inclusion) \mathcal{C}'-admissible sets. Thus, the following result follows from the previous one.

Theorem 10 Let \((X, \mathcal{D}, \mathcal{C})\) be a \mathcal{C}'-admissible CSP associated with a CAF \(F = (A, R, \mathcal{C})\). Each maximal (for set inclusion) set \(\{x_1, \ldots, x_k\}\) s.t. \(v(x_1) = 1\) and \((v(x_1), \ldots, v(x_n))\) is a solution of the CSP, is a \mathcal{C}'-preferred extension of \(F\).

Proof This follows from the definition of a \mathcal{C}'-preferred extension and Theorem 9.

5 Preference-based frameworks

Is is well acknowledged in argumentation literature that arguments may not have the same strength. For instance, arguments built from certain information are stronger than arguments built from uncertain information. Consequently, in [2] Dung’s framework was extended in such a way to take into account the strengths of arguments when evaluating them. The idea is to consider in addition to the attack relation, another binary relation \(\triangleright\) which represents preferences between arguments. This relation can be instantiated in different ways. Writing \(a \triangleright b\) means that \(a\) is at least as good as \(b\). Let \(\triangleright\) be the strict relation associated with \(\triangleright\). It is defined as follows: \(a \triangleright b\) if \(a \triangleright b\) and not \(b \triangleright a\). In Dung’s framework, an attack always succeeds (if the attacked argument is not defended). In preference-based frameworks, an attack may fail if the attacked argument is stronger than its attacker.
Definition 17 (PAF) A preference-based argumentation framework (PAF) is a tuple $\mathcal{F} = (\mathcal{A}, \mathcal{R}, \succeq)$ where $\mathcal{A}$ is a set of arguments, $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ is an attack relation and $\succeq$ is (partial or total) preorder on $\mathcal{A}$ ($\succeq \subseteq \mathcal{A} \times \mathcal{A}$).

The extensions of $\mathcal{F}$ (under any semantics) are those of the AF $(\mathcal{A}, \text{Def})$ where $(a, b) \in \text{Def}$ iff $(a, b) \in \mathcal{R}$ and not $(b \succeq a)$.

It is clear that each preference-based argumentation framework has a corresponding basic framework (in the sense of Definition 4). This latter is used for computing the extensions of the PAF under various semantics.

Property 7 Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}, \succeq)$ be a PAF and $\mathcal{T} = (\mathcal{A}, \text{Def})$ its corresponding AF. The extensions of $\mathcal{F}$ under semantics $x$ are exactly the extensions of $\mathcal{T}$ under the same semantics.

Proof This follows from the definition.

It is thus obvious that the extensions of a PAF $\mathcal{F} = (\mathcal{A}, \mathcal{R}, \succeq)$ under semantics $x$ are computed by the $x$-CSP that is associated with the basic framework $\mathcal{T} = (\mathcal{A}, \text{Def})$. For instance, the stable extensions of $\mathcal{F}$ are computed by the stable-CSP associated with $\mathcal{T}$, the preferred extensions of $\mathcal{F}$ are computed by the preferred-CSP associated with $\mathcal{T}$, and so on.

6 Related work

There are very few attempts in the literature for modeling argumentation frameworks as a CSP. To the best of our knowledge, the only work on the topic is the one done in [7]. In this work, the authors studied the problem of encoding weighted argumentation frameworks by semirings. In a weighted framework, attacks do not necessarily have the same weights. Thus, a weight (i.e., a value between 0 and 1) is associated with each attack between two arguments. When all the attacks have weight 1, the corresponding framework collapses with Dung’s abstract framework recalled in Section 3. The authors focused only on two semantics: stable and complete. In our paper, we proposed an alternative approach for computing the extensions of a basic argumentation framework ([19]) under all Dung’s semantics (i.e., even under admissible, preferred and grounded semantics). Moreover, we studied the same semantics for two other types of argumentation frameworks: constrained argumentation frameworks proposed in [13] and preference-based frameworks introduced in [2]. Finally, our approach for defining CSPs is simpler and more natural than that followed in [7]. Indeed, in [7], the authors used soft CSP whereas in our paper we used simple CSP.

Another interesting piece of work which is closer to our is that presented in [6]. Indeed, the authors encoded Dung’s semantics as a satisfiability problem (SAT). Besides, it was shown in [10] that SAT is a particular case of CSPs. Moreover, a mapping from SAT to CSP was given in the same paper. In our paper, we took advantage of that mapping and we presented different CSPs which encode Dung’s semantics not only for Dung’s framework, but also for constrained frameworks and preference-based ones.

Finally, a system called ASPARTIX implemented Dung’s semantics using answer set programming (ASP) [24]. ASPARTIX relies on a fixed disjunctive datalog program
which takes an argumentation framework as input, and uses the answer-set solver DLV for computing its extensions under a given semantics. This work is complementary to our especially since the ASP community established some links between ASP and CSP.

7 Conclusion

In this paper, we expressed the problem of computing the extensions of an argumentation framework under a given semantics as a CSP. We investigated three types of frameworks: Dung’s argumentation framework [19], its constrained version proposed in 13], and its extension with preferences [2]. For each of these frameworks, we proposed different CSPs which compute the extensions under various semantics, namely admissible, preferred, stable, complete and grounded. Such mappings are of great importance since they allow the use of the efficient solvers that were developed by CSP community. Thus, the efficiency of our CSPs depend on that of the solver that is chosen to solve them. The free-CSP has the form of 2-satisfiability (or 2SAT) problem. 2SAT consists of determining whether a collection of Boolean variables with constraints on pairs of variables can be assigned values satisfying all the constraints. This is particularly the case of free-CSP. It is acknowledged in SAT community that a 2SAT problem has a polynomial time solution [27]. In [25], polynomial time algorithms enumerating all the solutions are provided. It is worth mentioning that in the particular case of grounded semantics, there is an additional test of minimality that is required after computing the solutions of the corresponding CSP. This increases thus the complexity of computing the grounded extension of an argumentation framework. Consequently, this particular extension should be computed using existing algorithms in argumentation literature [1] and not by a CSP.

There are a number of ways to extend this work. One future direction consists of proposing the CSPs that return other semantics like semi-stable [9] and ideal [20]. Another idea consists of encoding weighted argumentation frameworks [22] as CSPs. In a weighted framework, attacks may not have the same importance. Such framework can be encoded by valued CSP in which constraints are associated with weights. Finally, we should study the kind of solvers that are more appropriate with our CSPs, and test their efficiency on real benchmarks.

References